AF]2172$

Attorney Docket No. YO998-467

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application

Applicant(s): Joseph L. Hellerstein
Docket No.: YO998-467
Serial No.: 09/285,639
Filing Date: April 2, 1999
Group: 2172
Examiner: Anh Ly

I hereby certify that this paper is being deposited on this date with the U.S. Postal Service as first class mail addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Signature: Denise A. Glaser Date: February 12, 2003

Title: Systems and Methods For Automated Navigation Between Dynamic Data With Dissimilar Structures

---

TRANSMITTAL OF APPEAL BRIEF

RECEIVED
MAR 06 2003
Technology Center 2100

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Submitted herewith are the following documents relating to the above-identified patent application:

(1) Appeal Brief in triplicate (original and two copies); and

(2) Copy of Notice of Appeal, filed on December 12, 2002, with copy of stamped return postcard indicating receipt of Notice by PTO on December 17, 2002.

Please charge **International Business Machines Corporation Deposit Account No. 50-0510** the amount of $320 to cover this submission under 37 CFR §1.17(c). In the event of non-payment or improper payment of a required fee, the Commissioner is authorized to charge or to credit **Deposit Account No. 50-0510** as required to correct the error. A duplicate copy of this letter and two copies of the Appeal Brief are enclosed.

Respectfully submitted,

William E. Lewis

Date: February 12, 2003

William E. Lewis
Reg. No. 39,274
Attorney for Applicant(s)
Ryan, Mason & Lewis, LLP
90 Forest Avenue
Locust Valley, NY 11560
(516) 759-2946

Attorney Docket No. YO998-467

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application

I hereby certify that this paper is being deposited on this date with the U.S. Postal Service as first class mail addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Signature: Denise A. Glaser Date: February 12, 2003

Applicant(s): Joseph L. Hellerstein
Docket No.: YO998-467
Serial No.: 09/285,639
Filing Date: April 2, 1999
Group: 2172
Examiner: Anh Ly

Title: Systems and Methods For Automated Navigation Between Dynamic Data With Dissimilar Structures

---

APPEAL BRIEF

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Applicant (hereinafter referred to as "Appellant") hereby appeals the final rejection of claims 1-27 of the above-identified application.

REAL PARTY IN INTEREST

The present application is assigned to International Business Machines Corporation, as evidenced by an assignment recorded May 7, 1999 in the U.S. Patent and Trademark Office at Reel 9946, Frame 0742. The assignee, International Business Machines Corporation, is the real party in interest.

RELATED APPEALS AND INTERFERENCES

There are no known related appeals and interferences.

STATUS OF CLAIMS

Claims 1-27 stand finally rejected under 35 U.S.C. §103(a). Claims 1-27 are appealed.

STATUS OF AMENDMENTS

There has been no amendment filed subsequent to the final rejection. However, a Response to Final Office Action was filed on December 12, 2002, along with the Notice of Appeal.

SUMMARY OF INVENTION

The present invention provides techniques that may aid in decision support applications by automatically selecting data relevant to an analysis. This may be accomplished by using the structure of the source dataset in combination with the content of the source element collection to identify the closest element collections within one or more target datasets (Specification, page 7, lines 23-27).

For example, the present invention, as recited in independent claim 1, defines a method of automating navigation between data with dissimilar structures including a source dataset containing one or more data elements and at least one target dataset containing one or more data elements. The method comprises the steps of: (i) determining at least one collection of data elements from the at least one target dataset that best matches a collection of data elements from the source dataset based on structures associated with the source dataset and the target dataset; and (ii) computing at least one distance metric between the at least one target collection and the source collection such that a user can select the at least one target collection given the at least one computed distance metric.

The invention may be implemented in a form which includes functional components as shown in FIG. 1A (Specification, page 12, line 12, through page 13, line 19). A first component is referred to as an inter-dataset navigation engine (IDNE) 140. The IDNE is invoked by analysis applications 110-1 through 110-N to automate the selection of related data. The IDNE makes use of another component referred to as dataset access services 130. The dataset access services component knows the accessible datasets and their structures, creates and manipulates collection descriptors, and provides access to elements within a dataset that are specified by a collection descriptor.

By way of example, in one embodiment, automated navigation according to the invention may be accomplished in the following manner. First, the IDNE 140 iterates across all target datasets to do the following: (a) use the structure of the source and target datasets to transform the source collection descriptor into a preliminary collection descriptor for the subset of the target dataset that

is closest to the source element collection; (b) construct the final collection descriptor by transforming the preliminary collection descriptor until it specifies a non-null subset of the target dataset; and (c) compute a distance metric representing how close the source element collection (or collection descriptor) is to the target element collection (or collection descriptor). The IDNE 140 then returns a list of triples including a name of the target dataset, a target collection descriptor, and a value of the distance metric for each target dataset (Specification, page 8, lines 9-19).

A flow diagram illustrating an automated navigation method according to an exemplary embodiment of the invention is shown in FIG. 2 (Specification, page 14, line 8, through page 15, line 13). Further, a flow diagram illustrating a technique for computing a target collection descriptor that best matches a source collection descriptor according to an exemplary embodiment of the invention is shown in FIG. 3 (Specification, page 15, line 14, through page 17, line 14). Still further, a flow diagram illustrating a technique for computing a distance metric according to an exemplary embodiment of the invention is shown in FIG. 4 (Specification, page 17, line 15, through page 18, line 10). Lastly, a diagram illustrating a graphical user interface for presenting exemplary results associated with automatic navigation according to an exemplary embodiment of the invention is shown in FIG. 5 (Specification, page 18, lines 11-19).

Accordingly, the present invention may provide automation for selecting datasets relevant to analysis tasks. Such automation is crucial to improving the productivity of decision support in systems management applications. The automation enabled by the invention provides value in many ways. For example, the invention makes the novice analyst more expert by providing a list of target datasets and collection descriptors that are closest to an element collection at hand (i.e., the source element collection). As a result, the novice focuses on the datasets that are most likely to be of interest in the analysis task. By way of further example, the invention makes expert analysis more productive. This is achieved by providing the target collection descriptor for each target dataset thereby enabling the construction of a system in which analysts need only click on a target dataset (or collection descriptor) in order to navigate to its associated element collection (Specification, page 9, line 24, through page 10, line 6).

Automated navigation according to the invention may be applied in many domains. For example, in analysis of manufacturing lines, measurement datasets exist for machines in the manufacturing line as well as for the interconnection of these machines. Automated navigation

according to the invention can aid with decision support for scheduling and planning based on this data. By way of a further example, in transportation systems, datasets exist for measurements taken by road sensors and traffic reports. Automated navigation according to the invention can aid in planning highway capacity over the entire network of roadways. It is to be understood that the above applications are merely exemplary in nature and not intended to limit the applicability of the invention. Furthermore, it is to be appreciated that automated navigation according to the invention can be accomplished centrally at a server or in a distributed manner amongst several smaller server machines (Specification, page 10, line 22, through page 11, line 4).

ISSUES PRESENTED FOR REVIEW

(I) Whether claims 1-4, 7, 12-15, 18, 23, 24 and 26 are unpatentable under 35 U.S.C. §103(a) over U.S. Patent No. 5,930,784 issued to Hendrickson (hereinafter "Hendrickson") in view of U.S. Patent No. 4,855,907 issued to Ferro, Jr. et al. (hereinafter "Ferro"); and

(II) Whether claims 5, 6, 8-11, 16, 17, 19-22, 25 and 27 are unpatentable under 35 U.S.C. §103(a) over Hendrickson in view of Ferro in further view of U.S. Patent No. 5,970,490 issued to Morgenstern et al. (hereinafter "Morgenstern").

GROUPING OF CLAIMS

Claims 1-27 stand or fall together.

ARGUMENT

Appellant incorporates by reference herein the disclosure of all previous responses filed in the present application, namely, responses dated: May 2, 2001; November 30, 2001; June 28, 2002; and December 12, 2002. Sections (I) and (II) to follow will respectively address issues (I) and (II) presented above. Section (III) will address the Examiner's comments offered in an Advisory Action dated December 31, 2002.

(I) With regard to the issue of whether claims 1-4, 7, 12-15, 18, 23, 24 and 26 are unpatentable under 35 U.S.C. §103(a) over Hendrickson in view of Ferro, the final Office Action contends that the cited combination discloses all of the claim limitations recited in the subject

claims. Appellant respectfully asserts that the combination of Hendrickson and Ferro fails to establish a prima facie case of obviousness under 35 U.S.C. §103(a), as specified in M.P.E.P. §2143.

As set forth therein, M.P.E.P. §2143 states that three requirements must be met to establish a prima facie case of obviousness. First, there must be some suggestion or motivation to combine reference teachings. Second, there must be a reasonable expectation of success. Third, the cited combination must teach or suggest all the claim limitations. While it is sufficient to show that a prima facie case of obviousness has not been established by showing that one of the requirements has not been met, Appellant respectfully believes that none of the requirements have been met.

As stated above, the present invention, for example as recited in independent claim 1, defines a method of automating navigation between data with dissimilar structures including a source dataset containing one or more data elements and at least one target dataset containing one or more data elements. The method comprises the steps of: (i) determining at least one collection of data elements from the at least one target dataset that best matches a collection of data elements from the source dataset based on structures associated with the source dataset and the target dataset; and (ii) computing at least one distance metric between the at least one target collection and the source collection such that a user can select the at least one target collection given the at least one computed distance metric. The independent claim had previously been amended in Appellant's response dated June 28, 2002 to further clarify that the determination step is based on structures associated with the source dataset and the target dataset. Independent claim 12 defines a similar apparatus-based invention, while independent claim 23 defines a similar article of manufacture-based invention. Independent claims 24 and 26 recite other embodiments of such automated navigation techniques. Each independent claim recites that the best match determination is based on structures associated with the source dataset and at least one target dataset.

On the other hand, Hendrickson discloses a method for locating related items in a geometric space which transforms relationships among items to geometric locations. More specifically, the method locates items in the geometric space so that the distance between items corresponds to the degree of relatedness. Hendrickson suggests that the method is especially beneficial for communicating databases with many items, and with non-regular relationship patterns. Examples of such databases with non-regular relationship patterns include databases containing items such as

scientific papers or patents, related by citations or keywords (see Abstract and column 3, lines 7 through 9 of Hendrickson).

However, as made clear in Hendrickson, the "relatedness" of any two items depends on substantive information associated with the items. That is, "non-regular relationship patterns" refers to substantive patterns. At column 3, lines 40 through 50, Hendrickson specifically discloses what "relatedness" means:

> Similarities between items can be based on many diverse characteristics of the items. For example, scientific papers can be similar if they contain common keywords. Alternatively, scientific papers can be similar if one paper cites the other paper, or if they both cite certain other papers. As another example, patents can be similar if they both cite the same other patent. Alternatively, they can be similar if they contain the same keywords, or if they share the same classification. Other characteristics can be used for assessing similarity, including geographic origin, time of origin, institutional origin, and authorship.

Ferro, on the other hand, discloses a method for moving a VSAM (virtual storage access method) base cluster to another DASD (direct access storage device) volume, while maintaining alternate indices into the cluster (see abstract and column 1 and 2 of Ferro).

First, Appellant asserts that no motivation or suggestion exists to combine Hendrickson and Ferro. For at least this reason, a prima facie case of obviousness has not been established. As is evident from the above summaries of the cited references, the two references perform different techniques, generating different results, in order to attempt to achieve different purposes. Hendrickson discloses a method for locating related items in a geometric space which transforms relationships among items to geometric locations. Ferro, on the other hand, discloses a technique for moving a VSAM base cluster to another DASD volume.

Appellant fails to see the motivation or suggestion to combine a geometric space-based item location system (Hendrickson) with a technique that deals with moving data in DASDs (Ferro). They are two completely unrelated concepts. As a result, Appellant strongly believes that one ordinarily skilled in the art would not look to a geometric space-based item location system to find inspiration to improve a system that attempts to move VSAM base clusters to another DASD volume, or vice versa.

Second, Appellant asserts that there is no reasonable expectation of success in achieving the present invention through a combination of Hendrickson and Ferro. For at least this reason, a prima facie case of obviousness has not been established. Despite the assertion in the final Office Action, Appellant does not believe that Hendrickson and Ferro are combinable since it is not clear how one would combine them. However, even if combined, for the sake of argument, they would not achieve a structure-based technique for automatically navigating between data with dissimilar structures including a source dataset containing one or more data elements and at least one target dataset containing one or more data elements, as the claimed invention provides.

Lastly, Appellant asserts that the combination of Hendrickson and Ferro fails to teach or suggest all of the claim limitations of independent claims 1, 12, 23, 24 and 26. For at least this reason, a prima facie case of obviousness has not been established. Again, assuming for the sake of argument that Hendrickson and Ferro could be properly combined, which for at least the reasons above it is believed that they can not be properly combined, the combination fails to teach or suggest all claim elements in independent claims 1, 12, 23, 24 and 26. The inventive steps (or operations) comprise determining at least one collection of data elements from at least one target dataset which best matches a collection of data elements from a source dataset based on structures associated with the source dataset and at least one target dataset; and then computing at least one distance metric between the target collection and the source collection such that the user can select the target collection.

In Appellant's previous response dated June 28, 2002, Appellant addressed how independent claims 1, 12, 23, 24 and 26 patentably distinguish over Hendrickson (since the Examiner's original grounds of rejection were based solely on Hendrickson). In the final Office Action, the Examiner added the Ferro reference, in combination with Hendrickson, to reject the independent claims. However, no where in the final Office Action (nor in the Advisory Action to be addressed below) does the Examiner address the claim language added in Appellant's previous response dated June 28, 2002 (i.e., the best match determination being <u>based on structures associated with the source dataset and at least one target dataset</u>), nor does the Examiner address the deficiencies previously pointed out with respect to Hendrickson regarding other claim limitations in the independent claims. The final Office Action appears to merely include the Ferro reference in combination with Hendrickson, and simply repeats the rejections raised in the previous non-final Office Action.

Nonetheless, as first addressed in Appellant's previous response dated December 17, 2002, Appellant will again address the deficiencies of the combination of Hendrickson and Ferro.

The cited combination, again assumed to be proper for argument sake, does not meet the limitations set out in independent claims 1, 12, 23, 24 and 26. As explained above, the claimed invention is directed to techniques for <u>automating navigation between data with dissimilar structures</u> which comprises determining at least one collection of data elements from the at least one target dataset that best matches a collection of data elements from the source dataset <u>based on structures associated with the source dataset and the target dataset</u>.

Hendrickson, as is evident, considers items based on the <u>substantive content in the items</u>, not <u>based on structures associated with the items</u>, as in the claimed invention. That is, for example, Hendrickson compares such substantive content as references, keywords, authors, etc., in determining relatedness between two papers such that the two papers may be locationally represented in a geometric space representing the items. Thus, Hendrickson does not account for <u>structural dissimilarity</u>, as does the claimed invention, but rather, accounts for <u>substantive similarity</u>. This is a fundamental difference between the two approaches.

Since Ferro has nothing to do with determining the relatedness of two items, Ferro does not make a best match determination whatsoever and, therefore, like Hendrickson, Ferro also <u>fails to disclose determining a best match based on structures associated with the source dataset and the target dataset</u>, as in the claimed invention.

Furthermore, Hendrickson and Ferro also fail to perform a distance metric computation <u>after</u> a determination of a best match, as recited by the claimed invention. Rather, Hendrickson uses a similarity computation to determine match or similarity. In fact, Hendrickson does not even determine a best match between items but merely determines how geometrically close two items should be placed in the geometric representation of the items. Ferro performs no best match determination. These are other fundamental differences between the Hendrickson/Ferro combination and independent claims 1, 12, 23, 24 and 26.

For a clear example of what type of data problem that the invention may provide a solution for with respect to dissimilar data structures, see the example provided in the context of QoS (quality of service) management at page 3, line 1, to page 4, line 27, of the present specification. This is

significantly different than the substantive similarity problem that Hendrickson attempts to address, and the DASD cluster movement problem that Ferro attempts to address.

Thus, the combination of Hendrickson and Ferro fails to disclose all of the limitations of independent claims 1, 12, 23, 24 and 26.

For at least the reasons given above, Appellant again respectfully requests withdrawal of the §103(a) rejections of independent claims 1, 12, 23, 24 and 26. Further, not only due to their respective dependence on such independent claims but also because such claims recite patentable subject matter in their own right, Appellant again respectfully requests withdrawal of the §103(a) rejections of dependent claims 2-4, 7, 13-15 and 18.

By way of example, the combination of Hendrickson and Ferro fails to teach or suggest the following claimed features: (i) wherein there is a plurality of target datasets from which respective best matching target collections are determined and respective distance metrics are computed such that the computed distance metrics are presented to the user in a ranked order (claims 2 and 13); (ii) wherein the presenting step further includes presenting the respective target collection to the user along with the respective computed distance metric (claims 3 and 14); (iii) wherein the presenting step further includes presenting a respective name associated with the target dataset to the user along with the respective target collection and the computed distance metric (claims 4 and 15); and (iv) wherein the data is organized in a multidimensional database and further wherein the determining step includes performing at least one drill-up operation on the target collection descriptor (claims 7 and 18).

(II) With regard to the issue of whether claims 5, 6, 8-11, 16, 17, 19-22, 25 and 27 are unpatentable under 35 U.S.C. §103(a) over Hendrickson in view of Ferro in further view of Morgenstern, the final Office Action contends that the cited combination discloses all of the claim limitations recited in the subject claims. Appellants assert that the combination of Hendrickson, Ferro and Morganstern fails to teach or suggest all of the claim limitations recited in the subject claims and, therefore, fails to establish a prima facie case of obviousness under 35 U.S.C. §103(a), as specified in M.P.E.P. §2143. Thus, Appellants again respectfully traverse the§103(a) rejection of claims 5, 6, 8-11, 16, 17, 19-22, 25 and 27 for at least the following reasons.

Appellants hereby re-allege and incorporate by reference the arguments relating to claims 1, 12, 23, 24 and 26 above in their entirety. Accordingly, due at least to the fact that claims 5, 6, 8-11, 16, 17, 19-22, 25 and 27 respectively depend from independent claims 1, 12, 24 and 26, it is believed that such dependent claims are allowable for at least the reasons identified above and, therefore, Appellants request withdrawal of the §103(a) rejections. That is, Morgenstern fails to remedy the deficiencies described above with respect to the Hendrickson/Ferro combination. Further, Appellants assert that claims 5, 6, 8-11, 16, 17, 19-22, 25 and 27 recite patentable subject matter in their own right.

By way of example, the combination of Hendrickson, Ferro and Morgenstern fails to teach or suggest the following claimed features: (i) wherein the source collection of data elements is specified by a source collection descriptor and the at least one target collection of data elements is specified by a target collection descriptor (claims 5 and 16); (ii) wherein the data is organized in a relational database and further wherein the determining step includes deleting at least one attribute associated with the target collection descriptor that is not present in the source collection descriptor (claims 6 and 17); (iii) wherein the determining step further includes the steps of: generating at least one preliminary target collection descriptor associated with the at least one target collection by transforming a source collection descriptor associated with the source collection; and removing constraints associated with the at least one preliminary target collection descriptor until a non-null element collection is obtained (claims 8, 19, 25 and 27); (iv) wherein the source collection of data elements is specified by a source collection descriptor and the at least one target collection of data elements is specified by a target collection descriptor and further wherein the computing step includes calculating the difference between constraints in the source collection descriptor and the target collection descriptor to compute the distance metric (claims 9 and 20); (v) wherein attributes of the constraints are weighted in accordance with their importance (claims 10 and 21); and (vi) wherein the distance metric is proportionally larger when the source and target collection descriptors differ by an attribute of a constraint that has a heavier weight associated therewith (claims 11 and 22).

(III) In the Advisory Action dated December 31, 2002, in response to Appellant's argument that there is no motivation or suggestion to combine Hendrickson and Ferro, the Examiner states that "Ferro teaches . . . moving datasets from source to target, that is, data is moved from one location to another from within a virtual storage access and its associated structure . . . and he also teaches a dataset permits accessing data sets efficiently in a single or multiple CPU environment . . . ." However, Appellant's fail to see how these statements provide the necessary motivation or suggestion to combine the cited references to obtain the claimed invention. Movement of data and access by multiple CPUs has nothing to do with the teachings of Hendrickson and, therefore, Appellant again reasserts that the cited combination is improper.

Further, in the Advisory Action, in response to Appellant's argument that the combination of Hendrickson and Ferro fails to teach or suggest all of the claimed limitations, the Examiner states that "Ferro . . . teaches a method of moving individual base clusters of VSAM from a source to a target DASD volume while maintaining the integrity of the data set and its associated structures . . . [and] how to distinguish between the source and target data sets . . . ." However, Appellant's again fail to see how these statements address all the deficiencies presented above and in Appellant's previous responses.

For at least the foregoing reasons, claims 1-27 are believed to be patentable over the cited references. As such, the application is asserted to be in condition for allowance, and favorable action is respectfully solicited.

Respectfully submitted,

William E. Lewis

Date: February 11, 2003

William E. Lewis
Attorney for Applicant(s)
Reg. No. 39,274
Ryan, Mason & Lewis, LLP
90 Forest Avenue
Locust Valley, NY 11560
(516) 759-2946

APPENDIX

1. A method of automating navigation between data with dissimilar structures including a source dataset containing one or more data elements and at least one target dataset containing one or more data elements, the method comprising the steps of:

determining at least one collection of data elements from the at least one target dataset that best matches a collection of data elements from the source dataset based on structures associated with the source dataset and the target dataset; and

computing at least one distance metric between the at least one target collection and the source collection such that a user can select the at least one target collection given the at least one computed distance metric.

2. The method of Claim 1, wherein there is a plurality of target datasets from which respective best matching target collections are determined and respective distance metrics are computed such that the computed distance metrics are presented to the user in a ranked order.

3. The method of Claim 2, wherein the presenting step further includes presenting the respective target collection to the user along with the respective computed distance metric.

4. The method of Claim 3, wherein the presenting step further includes presenting a respective name associated with the target dataset to the user along with the respective target collection and the computed distance metric.

5. The method of Claim 1, wherein the source collection of data elements is specified by a source collection descriptor and the at least one target collection of data elements is specified by a target collection descriptor.

6. The method of Claim 5, wherein the data is organized in a relational database and further wherein the determining step includes deleting at least one attribute associated with the target collection descriptor that is not present in the source collection descriptor.

7. The method of Claim 5, wherein the data is organized in a multidimensional database and further wherein the determining step includes performing at least one drill-up operation on the target collection descriptor.

8. The method of Claim 1, wherein the determining step further includes the steps of:

generating at least one preliminary target collection descriptor associated with the at least one target collection by transforming a source collection descriptor associated with the source collection; and

removing constraints associated with the at least one preliminary target collection descriptor until a non-null element collection is obtained.

9. The method of Claim 1, wherein the source collection of data elements is specified by a source collection descriptor and the at least one target collection of data elements is specified by a target collection descriptor and further wherein the computing step includes calculating the difference between constraints in the source collection descriptor and the target collection descriptor to compute the distance metric.

10. The method of Claim 9, wherein attributes of the constraints are weighted in accordance with their importance.

11. The method of Claim 10, wherein the distance metric is proportionally larger when the source and target collection descriptors differ by an attribute of a constraint that has a heavier weight associated therewith.

12. Apparatus for automating navigation between data with dissimilar structures including a source dataset containing one or more data elements and at least one target dataset containing one or more data elements, the apparatus comprising:

at least one processor operable to determine at least one collection of data elements from the at least one target dataset that best matches a collection of data elements from the source dataset based on structures associated with the source dataset and the target dataset, and to compute at least

one distance metric between the at least one target collection and the source collection such that a user can select the at least one target collection given the at least one computed distance metric; and

a memory coupled to the at least one processor for storing the at least one target dataset.

13. The apparatus of Claim 12, wherein there is a plurality of target datasets from which respective best matching target collections are determined and respective distance metrics are computed such that the computed distance metrics are presented to the user in a ranked order.

14. The apparatus of Claim 13, wherein the at least one processor is further operable to present the respective target collection to the user along with the respective computed distance metric.

15. The apparatus of Claim 14, wherein the at least one processor is further operable to present a respective name associated with the target dataset to the user along with the respective target collection and the computed distance metric.

16. The apparatus of Claim 12, wherein the source collection of data elements is specified by a source collection descriptor and the at least one target collection of data elements is specified by a target collection descriptor.

17. The apparatus of Claim 16, wherein the data is organized in a relational database and further wherein the at least one processor is operable to perform the determining operation by deleting at least one attribute associated with the target collection descriptor that is not present in the source collection descriptor.

18. The apparatus of Claim 16, wherein the data is organized in a multidimensional database and further wherein the at least one processor is operable to perform the determining operation by performing at least one drill-up operation on the target collection descriptor.

19. The apparatus of Claim 12, wherein the at least one processor is further operable to perform the determining operation by generating at least one preliminary target collection descriptor associated with the at least one target collection by transforming a source collection descriptor associated with the source collection, and removing constraints associated with the at least one preliminary target collection descriptor until a non-null element collection is obtained.

20. The apparatus of Claim 12, wherein the source collection of data elements is specified by a source collection descriptor and the at least one target collection of data elements is specified by a target collection descriptor and further wherein the at least one processor is operable to perform the computing operation by calculating the difference between constraints in the source collection descriptor and the target collection descriptor to compute the distance metric.

21. The apparatus of Claim 20, wherein attributes of the constraints are weighted in accordance with their importance.

22. The apparatus of Claim 21, wherein the distance metric is proportionally larger when the source and target collection descriptors differ by an attribute of a constraint that has a heavier weight associated therewith.

23. An article of manufacture for automating navigation between data with dissimilar structures including a source dataset containing one or more data elements and at least one target dataset containing one or more data elements, comprising a machine readable medium containing one or more programs which when executed implement the steps of:

determining at least one collection of data elements from the at least one target dataset that best matches a collection of data elements from the source dataset based on structures associated with the source dataset and the target dataset; and

computing at least one distance metric between the at least one target collection and the source collection such that a user can select the at least one target collection given the at least one computed distance metric.

24. A computer-based method of automatically navigating between data with dissimilar structures including a source dataset containing one or more data elements and a plurality of target datasets respectively containing one or more data elements, the method comprising the steps of:

determining one or more collections of data elements from the plurality of target datasets that best match a collection of data elements from the source dataset, the determination being based on the structures associated with the source dataset and the plurality of target datasets; and

computing one or more distance metrics between the one or more target collections and the source collection.

25. The method of Claim 24, wherein the determining step further comprises the steps of:

generating one or more preliminary target collection descriptors associated with the one or more target collections by transforming a source collection descriptor associated with the source collection; and

removing constraints associated with the one or more preliminary target collection descriptors until a non-null element collection is obtained.

26. Apparatus for automatically navigating between data with dissimilar structures including a source dataset containing one or more data elements and a plurality of target datasets respectively containing one or more data elements, the apparatus comprising:

means for determining one or more collections of data elements from the plurality of target datasets that best match a collection of data elements from the source dataset, the determination being based on the structures associated with the source dataset and the plurality of target datasets; and

means for computing one or more distance metrics between the one or more target collections and the source collection.

27. The apparatus of Claim 26, wherein the determining means further comprises:

means for generating one or more preliminary target collection descriptors associated with the one or more target collections by transforming a source collection descriptor associated with the source collection; and

means for removing constraints associated with the one or more preliminary target collection descriptors until a non-null element collection is obtained.